VISIT…

LANZAROTE Caliente.COM

Материалы X Всероссийского семинара "Нейроинформатика и ее приложения", Красноярск: КГТУ, 2002. 185с. – с.157-160.

ОБ ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ ОДНОГО МЕТОДА ПОСТРОЕНИЯ ОТКАЗОУСТОЙЧИВЫХ НЕЙРОСЕТЕЙ

В.Г. Царегородцев

www.NeuroPro.ru tsar@neuropro.ru

Реализован и опробован алгоритм обучения, временно прекращающий адаптацию наиболее значимых синапсов при обучении нейросети обратного распространения. Полученные нейросети обладают большей устойчивостью к отказам элементов и лучшей обобщающей способностью, что подтверждается известными результатами.

1. Идеи и алгоритмы

Исследуется предложенный в [1] способ увеличения отказоустойчивости нейронной сети. Кратко, идея состоит из следующих пунктов:

1. Отказоустойчивость реализуется в ходе обучения, способом, внешним к стандартному алгоритму обучения и не зависящим от него.
2. Для увеличения отказоустойчивости необходимо не допускать слишком большого роста значимости (информативности, релевантности) отдельных синапсов сети.
3. Для реализации указанного в п.2 критерия параллельно с процессом обучения осуществляется мониторинг значений показателей значимости синапсов и заданное число наиболее значимых синапсов временно выводится из обучения.

Естественно, идею можно расширять для получения устойчивости к отказам нейронов скрытых слоев нейросети (путем оценки значимостей выходных сигналов нейронов и запрещения обучения выходных синапсов наиболее значимых нейронов) и для снижения чувствительности нейросетевого решения к отдельным входным сигналам сети (путем оценки значимостей входных сигналов и вывода из обучения синапсов, по которым значимые входные сигналы проходят на нейроны скрытого слоя).

Достоинствами такого подхода является независимость от топологии нейросети, алгоритма обучения и способа вычисления значимостей сигналов/элементов нейросети.

Хотя простой вывод отдельных наиболее значимых элементов из обучения в общем случае не гарантирует снижения их значимости относительно других элементов при дальнейшем обучении, избыточный рост значимостей отдельных элементов (как это бывает при обычном обучении нейросети стандартными алгоритмами) все же в итоге не наблюдается – это демонстрируется как в оригинальной работе [1], так и здесь.

Однако, базовая идея все же накладывает некоторые ограничения. Постоянное изменение набора адаптивных параметров сети при вводе/выводе синапсов в обучение требует использования для обучения сети обычного алгоритма наискорейшего градиентного спуска – применение более быстрых алгоритмов (метод сопряженных градиентов, квазиньютоновские алгоритмы) невозможно. Также оценка значимостей параметров одновременно с обучением (в пределе – оценка значимости каждого адаптивного параметра после каждого шага градиентного спуска) требует собственных временных затрат и, соответственно, необходимости в настройке баланса между скоростью общей работы и точностью оценки значимостей (путем выбора конкретного метода оценки значимостей и определения числа итераций градиентного спуска, разделяющих моменты перевычисления показателей значимости).

Поскольку давно известен факт, что абсолютная величина значения параметра, например, модуль веса синапса, плохо оценивает реальную значимость параметра (чувствительность решения к изменению значения этого параметра), см., например, [2], необходимо применять более сложные способы – оценивание изменения значения целевой функции или выходных сигналов сети на выборке описанными в [2-7] методами первого или второго порядка. Точное же вычисление реальных значимостей

элементов (например, точное вычисление изменения значения целевой функции на выборке при элементарной модификации нейросети – [3]) требует слишком уж больших суммарных затрат для всех элементов нейросети и здесь неприменимо.

2. Результаты экспериментов

Эксперименты проводились для задачи классификации с учителем – задачи определения направления циклического сдвига восьмибитной последовательности (более подробное описание задачи см. в [7]).

Обучались нейронные сети с топологией 16-5-2 (число входных сигналов, число нейронов в скрытом слое, число выходных нейронов, всего в сети 97 синапсов) с использованием специальной целевой функции вида "расстояние до множества" [4,6]. Обучение велось до момента правильного распознавания класса для всех примеров обучающей выборки – обучение до достижения нулевого значения суммарной целевой функции приводило в этой задаче к возникновению эффекта переобучения нейросетей.

Использовались показатели значимости первого порядка, оценивающие изменение выходных сигналов нейросети [7]. Значения показателей значимости всех синапсов сети пересчитывались после каждого шага градиентного спуска. Градиентный алгоритм обучения на очередной итерации не модифицировал 10 наиболее значимых синапсов сети – т.е. на каждом шаге обучения коррекция 10 наиболее значимых синапсов не производилась. Столь частое (на каждом шаге обучения) обновление показателей значимости вызвано возможностью значительного их изменения за малое число шагов – пример динамики показателей значимости первого порядка см. в [8].

Рис.1, Таблица 1 дают среднее число неправильных ответов на обучающей выборке для двух нейросетей – полученной при обычном обучении и путем применения описанной в предыдущем абзаце процедуры. Усреднение велось по всем возможным вариантам одновременного отказа (обращения веса синапса в нуль) заданного числа синапсов. Специально обученная сеть более отказоустойчива.

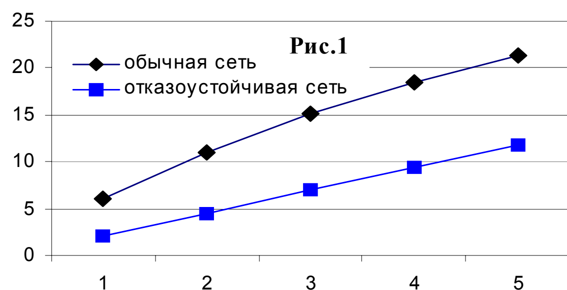


Таблица 1	Число отказов синапсов				
Нейросеть	1	2	3	4	5
обычная	6.02	11.03	15.13	18.53	21.4
отказоустойчивая	2.11	4.47	6.94	9.39	11.75

Среднее число неправильных ответов, в % от объема выборки, на обучающей выборке для всех возможных комбинаций отказов от 1 до 5 синапсов для отказоустойчивой и обычной сетей.

На Рис.2 показано распределения значений весов синапсов обычной и полученной отказоустойчивой нейросетей, на Рис. 3 – распределения величин значимостей синапсов для этих нейросетей. По оси ординат на Рис.2-3 дано число синапсов (которое из-за наличия в сети 97 синапсов близко к значениям частот в процентах). Отказоустойчивые сети имеют меньшее число синапсов с малыми весами (лежащими в диапазоне [-0.1,0.1] начальной инициализации весов) – из-за того, что "замораживание" значимых синапсов потребовало адаптации иных синапсов, которые при обычных условиях могли бы адаптации не подвергнуться и остаться практически с исходными значениями весов. Рис.3 подтверждает это утверждение об адаптации изначально избыточных синапсов и "осреднении" значимостей синапсов в отказоустойчивой сети – в отказоустойчивой сети по сравнению с обычной меньше синапсов с минимальной значимостью, но больше синапсов с средней значимостью и значимостью ниже среднего. Максимальные же значения значимостей синапсов в

отказоустойчивой сети меньше, чем в обученной стандартным способом, причем виден явный пик частоты для максимальных значимостей, вызванный алгоритмом ограничения роста значимости.

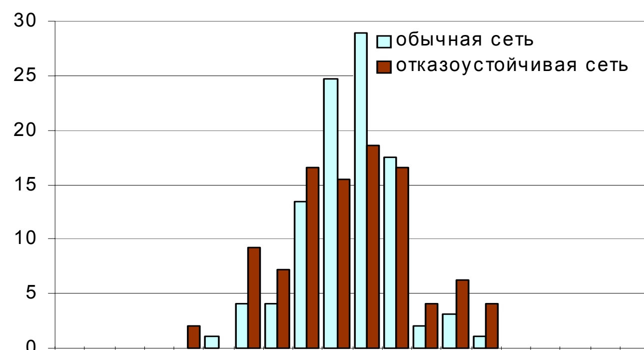


Рис.2. Распределение значений весов синапсов в диапазоне [-1,1] после обучения сетей.

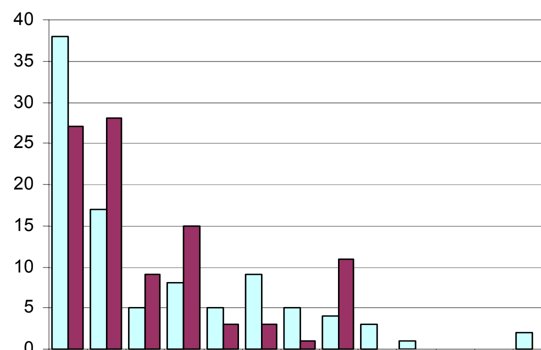


Рис.3. Распределение значений показателей значимости синапсов после обучения сетей.

Хорошо известен факт, что использование при обучении сети дополнительного критерия вторичной оптимизации (обычно регуляризующего критерия) часто увеличивает обобщающие способности нейросети и скорость сходимости процесса обучения. Стандартно критерий вторичной оптимизации учитывается в суммарной целевой функции наряду с первичным критерием качества работы нейросети. Однако, построение критерия вторичной оптимизации в виде процедуры, внешней по отношению к алгоритму обучения, также часто приводит к значительному положительному эффекту. Можно надеяться, что реализованный здесь "квазикритерий" вторичной оптимизации, не допускающий роста значимостей синапсов, также будет вдобавок увеличивать и обобщающую способность отказоустойчивых нейросетей.

Для рассмотренной задачи и данного способа достижения отказоустойчивости действительно наблюдается увеличение обобщающих свойств отказоустойчивых сетей. На Рис.4 показаны величины ошибок обобщения, полученные в процессе скользящего экзамена по обучающей выборке. Обучались копии 20 начальных нейросетей, сначала обычным градиентным обучением, затем с применением описанного способа увеличения отказоустойчивости. Точками на Рис.4 даны величины ошибок обобщения, горизонтальным маркером – средние значения ошибок обобщения.

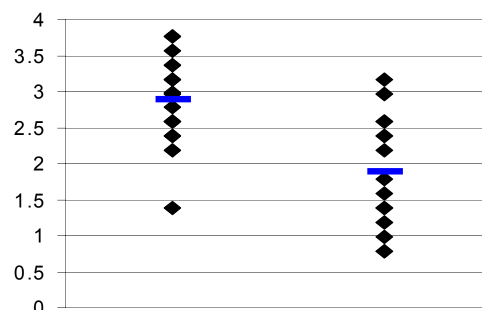


Рис.4. Ошибка обобщения, в % от объема выборки, для 20 обученных нейросетей. Слева – обычные сети, справа – отказоустойчивые.

Интересен вопрос об устойчивости наблюдаемых эффектов и их зависимости от настроек алгоритма (числа выводимых из обучения наиболее значимых синапсов, числа шагов обучения между перерассчетами значимостей). Естественно, какие-либо тенденции могут существовать только "в среднем" – для набора однотипных нейросетей, различающихся начальными наборами значений синапсов. Это затрудняет применимость отдельных способов точного анализа (например, анализа средней или максимальной ошибки решения путем полного перебора вариантов отказов заданного типа – как было сделано выше при получении отраженных на Рис.1, Таблице 1 результатов). Для иллюстрации гипотезы о существовании только "средних" тенденций дополним Рис.1 данными по отказоустойчивости сетей, полученных при обучении исходной сети с выводом из обучения двух, четырех, шести и восьми наиболее значимых синапсов соответственно. Полученные результаты даны на Рис.5 и говорят,

что алгоритм обеспечения отказоустойчивости, по крайней мере для использованной базовой необученной сети и пяти разных настроек числа выводимых из обучения синапсов (2,4,6,8,10 синапсов) обеспечивает лучшую отказоустойчивость по сравнению с обычным обучением. Хоть монотонной зависимости уровня отказоустойчивости от числа выведенных из обучения наиболее значимых синапсов для одной сети не наблюдается, монотонность может проявиться в среднем для нескольких сетей.

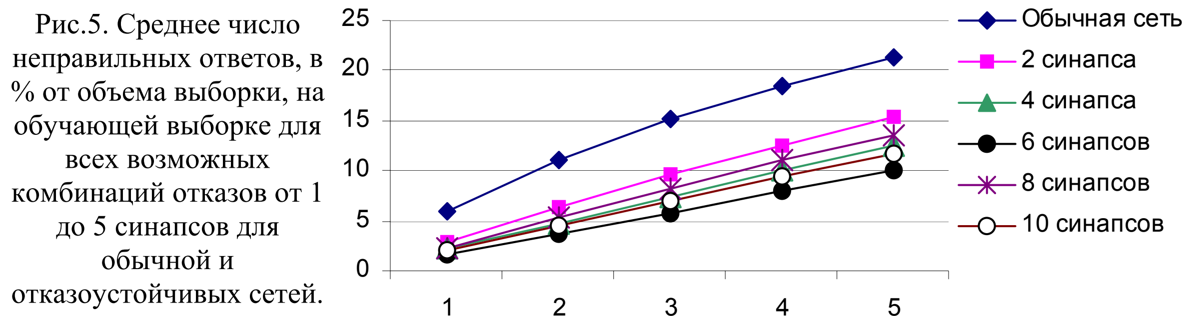


Рис.5. Среднее число неправильных ответов, в % от объема выборки, на обучающей выборке для всех возможных комбинаций отказов от 1 до 5 синапсов для обычной и отказоустойчивых сетей.

3. Заключительные замечания

Проблема построения отказоустойчивых нейросетей в принципе достаточно широко исследована в спецлитературе. Автор при изучении этой темы и проведении экспериментов сформировал следующее собственное отношение к проблеме.

Наиболее перспективными и адекватными представляются алгоритмы, обеспечивающие достижение отказоустойчивости в ходе обучения нейросети. К такому классу алгоритмов принадлежат исследованный здесь метод [1] и метод учета шумовых возмущений синапсов при обучении сети в целевой функции [9]. Способ же [5] неоднократного дообучения сети с шумовым возмущением между этапами дообучения не всегда эффективен, поскольку требует внесения шума со значительной амплитудой, что может привести к большим (по модулю) весам синапсов и, соответственно, активации отдельных нейронов только в областях насыщения нелинейной функции, т.е. к реальному ограничению информационной емкости сети и недостижению максимальной отказоустойчивости. Конечно, существуют и иные методы, но объем статьи не позволяет на них остановиться даже путем простого перечисления работ.

Литература

1. Kwon O.J., Bang S.Y. *A design method of fault tolerant neural networks* / Proc. ICONIP'94, Seoul, Korea. 1994. Vol.1. – pp. 396-400.
2. Le Cun Y., Denker J.S., Solla S.A. *Optimal Brain Damage* / Advances in Neural Information Processing Systems 2 (1989). Morgan Kaufmann, 1990. – pp.598-605.
3. Mozer M.C., Smolensky P. *Skeletonization: a technique for trimming the fat from a network via relevance assessment* / Advances in Neural Information Processing Systems 1 (1988). Morgan Kaufmann, 1989. – pp.107-115.
4. Миркес Е.М. *Нейрокомпьютер: проект стандарта*. Новосибирск: Наука, 1999. - 337с.
5. Горбань А.Н. *Обучение нейронных сетей*. М.: ПараГраф, 1990. - 160с.
6. Горбань А.Н., Россиев Д.А. *Нейронные сети на персональном компьютере*. Новосибирск: Наука, 1996. - 276с.
7. Царегородцев В.Г. *Простейший способ вычисления показателей значимости первого порядка для сетей обратного распространения* // Материалы X Всеросс. семинара "Нейроинформатика и ее приложения", Красноярск: КГТУ, 2002. 185с. – с.153-156.
8. Engelbrecht A.P., Cloete I. *A sensitivity analisys algorithm for pruning feedforward neural networks* / Proc. IEEE ICNN 1996, Washington, DC, USA, Vol.2. – pp.1274-1278.
9. Murray A.F., Edwards P.J. *Synaptic weight noise during MLP learning enhances fault-tolerance, generalization and learning trajectory* / Advances in Neural Information Processing Systems 5 (1992). Morgan Kaufmann, 1993. – pp.491-498.